DOCUMENT RESUME

ED 111 863 TM 004 830

AUTHOR Loret, Peter G.
TITLE Implementing, Evaluating, and Using a Statewide Assessment Program: Logistics and Contracted Services.
PUB DATE [Apr 75]
NOTE 13p.; Paper presented at the Annual Meeting of the National Council on Measurement in Education (Washington, D.C., March 31-April 2, 1975)

EDRS PRICE MF-$0.76 HC-$1.58 Plus Postage
DESCRIPTORS *Educational Assessment; Elementary Education; Factor Analysis; *Grade 4; Program Development; Reading Skills; *Reading Tests; Scoring; *State Programs; Statistical Analysis; Test Construction; *Testing
IDENTIFIERS *Oregon Statewide Assessment Program

ABSTRACT

Under contract with the Oregon State Department of Education, Educational Testing Service (ETS) assisted in the preparation of a fourth grade reading assessment test based on, or modified from, Instructional Objectives Exchange materials. ETS produced test books, answer sheets, and all required ancillary materials necessary for test distribution and administration to a sample of 104 schools, previously selected on the basis of geographic region, district wealth, and district size. Rights, Wrongs, and "Don't Know" pupil scores were reported to schools and districts for 25 specific reading objectives and 4 grouped objectives (word attack, vocabulary, comprehension, and application). Summary statistics for schools were also prepared and distributed. A special vocabulary self-report section/score was included in both reports. At the state level, a number of item analyses were performed, and a series of weight frequency distributions prepared. A total of 4,127 pupils at 102 schools actually took the test with 299 pupils reported as absent. (Author)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED111863

Implementing, Evaluating and Using a Statewide Assessment Program: Logistics and Contracted Services[1]

by
Peter G. Loret
Educational Testing Service
Berkeley, California

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

The Sample

At the time of the award of the contract for the Oregon Statewide Assessment to Educational Testing Service, the sample, consisting of 104 independently drawn and randomly selected Oregon schools, had already been identified under a separate contract between the Oregon State Department of Education and the Research Triangle Institute (RTI)[2]. The Oregon State Department of Education provided ETS with a list of school names and addresses, together with initial estimates of fourth grade enrollment, an Oregon-assigned identification number (SDE number), and a special RTI identification number specifying sample stratification variables (geographic region, district wealth, district size). The Oregon State Department of Education also provided the names and addresses of school principals, the Intermediate Education District (IED/County) Coordinators, and the district superintendents who were to receive score reports. All direct contacts with schools, districts, and IED/County units regarding participation, score reporting, and similar matters were handled directly by the Oregon State Department of Education.

Initial enrollment estimates for schools were later modified to reflect actual enrollment, and the quantities of materials adjusted to reflect these data.

---

[1]Prepared for presentation at the meetings of the National Council on Measurement in Education, Washington, D.C., April 2, 1975.

[2]Bayless, David L. and Weber, Jilene: Sample Design of Oregon Assessment of Fourth Grade Reading Performance in the Spring of 1973, Final Report 22U-951, Research Triangle Institute - Center for Educational Research and Evaluation, 19 pp.

TM004 830

The initial Oregon sample consisted of 104 schools, containing an estimated 4,544 fourth-grade pupils, located in 29 different Oregon IED/ County units. Revised figures, based on actual enrollment, increased the sample size to 4,584 pupils.

The Test

The initial pool of test items (questions) which constituted the source of the actual instrument administered to Oregon fourth-graders was originally obtained by the Oregon State Department of Education from the Instructional Objectives Exchange (IOX). Test items were initially selected and revised by the staff of the Oregon State Department of Education in order to provide measurement of twenty-five (25) specific reading objectives and four grouped objectives or "skills" areas (Word Attack, Vocabulary, Comprehension, and Application).

The test items were systematically reviewed by ETS staff to confirm their relevance to the stated objectives, to minimize bias and possible offensiveness to minorities, and to judge their technical quality in terms of content and face validity. ETS staff suggested item revisions and modifications and met with the Oregon State Department of Education staff and an outside reading consultant to finalize the content and format of the test.

In addition to the ninety-six (96) "standard" test items selected for inclusion in the test, a special vacabulary self-report section was also developed and included. This section consisted of a list of vocabulary items, drawn from the test and test directions, for which the examinee was to indicate whether the word was understood (i.e., "clear" or "not clear"). Following approval of the test, ETS designed a machine-scorable answer sheet for use on its NCS equipment; the answer sheet contained both the

actual questions and the answers (responses) for the first forty test items, and provided response spaces for the last fifty-six items (contained in a separately printed test book.) The answer sheet also contained the check-list of thirty-eight vocabulary items.

Following test production and printing, but prior to actual test administration, three items were identified as faulty (items 13, 54, and 55) and were deleted from all subsequent scoring, processing, reporting and analysis.

Preparation of Materials and Shipment

As specified in its contract, ETS prepared and submitted a Basic Procedural Design Manual containing drafts of the form letters, manuals, test books, answer sheets, and ancillary materials to be used in the administration of the program. Following review, revision, and final approval, appropriate quantities of each item were printed for packing and distribution. With the exception of one large shipment to Multnomah County, all packages were sent to IED/County Coordinators via United Parcel Service (UPS). Return shipments were automatically picked up by UPS on specified dates (April 30 for all IEDs testing one or two schools and May 3 for all IEDs testing three or more schools). UPS service attempts to pick up return shipment(s) on three successive days, beginning with the scheduled date, and was to notify ETS if return shipments were unavailable. All packages were returned as scheduled and without additional follow-up.

Training Sessions (Workshops)

During the week of April 9-12, 1974, the staff of the Oregon State Department of Education, assisted by the ETS project director, conducted a

series of one-day training session/workshops in Wilsonville, Sutherlin, and Pendleton. These meetings were attended by IED/County, district and school personnel, and provided general background to participants on the Oregon Statewide Assessment Program. The sessions also included a detailed review and discussion of the administrative procedures to be utilized in the program. Sample materials were distributed, reviewed, and discussed by the participants. Copies of the materials, specially prepared for the workshops, were made available to the participants.

Test Administration

The test was to be administered to fourth-grade pupils in the selected schools on April 24, 1974. Pupils exempted from taking the test included those officially classified as blind, deaf, physically handicapped (unable to manipulate pencil and paper), trainable mentally retarded, or emotionally disturbed (to the extent that testing them would have tended to disturb other pupils being tested). Schools unable to test on the specified date were asked to administer the instrument during the week of April 22nd. In addition to administering the test to pupils present on the test date, test administrators (teachers) were asked to complete an answer sheet for those pupils absent on the date of testing, and to provide an indication that the pupil had been absent on the test date by marking the first entry in the "For Test Administrator Use Only" area on the answer sheet. Make-up administrations were not to be scheduled. (4,414 pupils at 102 schools actually took the test; there were 306 absentees and 10 invalid cases, providing usable data for 4,098 pupils for further analyses.)

Post-Administration Activities by Schools and IED/County Coordinators

Following the actual test administration, the teachers (test admin-

istrators) were asked to supply individual pupil data on the front of each pupil's answer sheet. These data included information on whether the pupil was officially classified as handicapped, his/her racial or national origin, participation in federally funded programs, ease of providing information on participation in a migrant pre-school program, bilinguality (and whether, if pupil was bilingual, this constituted a learning handicapped), whether the pupil had been specifically diagnosed as needing corrective or remedial work in reading, and whether he/she was participating in a corrective or remedial reading program.

School coordinators were requested to verify that all students had gridded properly, and to be sure that the teachers had supplied all individual pupil data. After completing a School Coordinator's Report Form, all materials were returned to appropriate IED/County Coordinator. IED/County Coordinators were asked to do additional editing prior to the return of materials to ETS.

As a condition of the contract with the Oregon State Department of Education, IED units were reimbursed (only upon specific request indicated on the IED/County Coordinator's Report Form) for expenses generated in the administration of the assessment program. The maximum amount of reimbursement was not to exceed $200 plus 25 cents per document edited, up to a maximum of $400. Nine IEDs did not request reimbursement; two others did not return the Report Form.

Materials Receipt; Accuracy of IED Editing

When all test materials had been returned by the IED/County Coordinators, test books were checked and verified against the materials originally shipped. A written report of missing materials was transmitted to the Oregon State Department of Education for follow-up and resolution.

As part of its work, ETS was asked to analyze the accuracy of IED/ County Coordinator editing, and to compare the results for those who had been reimbursed with those who had not requested payment. A comparison of those IED Coordinators requesting payment of an honorarium (N = 18) with those not requesting payment (N = 9) appears to suggest, at least on a proportional basis, that those who did not request payment performed the editing function less accurately than those requesting payment. The errors that did exist seemed to cluster in a single school, rather than within an individual IED.

Scoring, Processing, and Reporting -: Schools and Districts

All answer sheets were optically scanned, transcribed, and edited in accordance with editing rules and procedures approved by the Oregon State Department of Edcuation.

During the course of machine processing, a sample of approximately 45 quality control cases (every 100th record) was identified. These cases were hand-processed, and the results compared with the computer output. All reports were verified as correct prior to release of actual score reports to participating schools and districts.

Two output reports were designed by ETS, reviewed, revised, and approved by the Oregon State Department of Education. The Student Report consisted of an alphabetic roster of pupil names and, for each individual, the number of responses marked Right, Wrong, or Don't Know for the twenty-five objectives and the four skills areas. In addition, the number of Word Identification items marked "Clear" or "Not Clear" was shown. Absentees were identified by the legend "Absentee", and any pupil whose test contained four or fewer responses was identified with the legend "Not Attempted".

The School Report displayed the twenty-five objectives, the four "grouped objectives" (i.e., skills areas), and the Word Identification score, and provided, for each heading, the number of questions, the mean number or Right responses per pupil and standard deviation for that school.

The Student and School Reports, together with a document entitled "Interpretive Guide - Pupil and School Report" were mailed to participating schools and district superintendents.

Based on an ETS magnetic tape containing all data transcribed in the course of the program, the Oregon State Department of Education contracted for the conduct of an independent verification of three aspects of ETS processing[3]. For a sample of 459 answer sheets selected by the Oregon State Department of Education, the data were coded and compared with ETS data. In addition, RTI, SDE, and ETS codes were matched and verified. Finally, the scoring of correct responses, by objectives and skills, was checked. No discrepancies were found in the latter two cases. In performance of the scanning function, 27 discrepancies were found in almost 80,000 characters compared (a discrepancy rate of 0.0338%); only one of the twenty-seven discrepancies was a direct ETS error not attributable to the scanning technique.

Other Analyses

Following the reporting of scores, a number of data analyses were undertaken. These included the item analysis of the test questions both for a random statewide sample of the 4,098 students tested, and for subsamples based on racial subgroups within the population. Intercorrelations were computed and a factor analysis performed.

[3]Memorandum from Nileen Hunt, Research Triangle Institute, 7-1-74, Verification of the 1973-74 Oregon Statewide Assessment Data Tape, 3 pages.

Weighted frequency distributions of scores for each reading objective and for the four "skills" areas were provided on each of the following groups:

a) the total State sample

b) by geographic region

c) by district wealth

d) by district size

e) by district size within geographic region

f) for the Metropolitan district

g) by two racial subgroups

Summary of Results

The test for the Oregon Statewide Assessment was designed to measure pupil performance on twenty-five reading objectives as well as four basic skills: Word Attack, Vocabulary, Comprehension, and Application. These skills are usually considered to be relatively independent components of the reading process, and most reading curricula involve specific instruction in each skill area. However, their measurement at the elementary school level typically reflects a high degree of interdependence among the skills. This interdependence is reflected in the reading test used in the Oregon program.

The correlations among the four basic skill areas are shown in table 1, as are the reliability estimates (KR 20) for each of the four areas of the test.

Insert Table 1 about here

Pupil performance is summarized in Table 2. Because each skill area contains a different number of test items, the average number of correct responses for each skill area is also expressed as a percentage of the number of items.

Most standardized tests of reading are designed to have an average

difficulty level of 55 to 65 percent of the items. This level is selected so that the instrument will yield reliable measurement throughout the entire range of reading ability. In the case of the Oregon data, the performance of pupils on the Word Attack, Vocabulary, and Comprehension sections significantly exceeds this level. Thus, only the Application Section of the test has an adequate difficulty level to reliably assess the reading skills of the better readers.

Insert Table 2 about here

It is reasonable to infer that, due to the characteristics of the test itself, most pupils found the test relatively easy and that the "bright" pupils, therefore, were unable to demonstrate the full extent of their reading skills. While this lack of discriminatory power at the top of the ability range does not seriously affect the validity of the test results for heterogeneous groups of pupils, this limitation must be taken into consideration when interpreting group data and the factor analytic results presented in Table 3.

Insert Table 3 about here

The factor analytic structure of the test (the statistical analysis of inter-relationships of the objectives) indicates a strong "Word Attack" factor (defined by Objectives 1 through 9), a strong "Comprehension" factor (defined by objectives 12, 13, and 15 through 21) and a weak "Application" Factor (defined primarily by objectives 23 through 25, which consist of a total of only four test items). The complex structure of the Vocabulary objectives (as indicated on Factors 1 and 2) may be attributed to the fact that the Vocabulary items were easy, i.e., most pupils earned perfect scores on the Vocabulary objectives. In this situation, the correlation of objectives 10 and 11 with all other objectives would be expected to be

essentially of the same magnitude. Factor 2 clearly indicates a high degree of interdependence between Comprehension and Application Objectives.

Conclusion

This has been an attempt to report the operational highlights of the first or pilot year of Oregon's Statewide Assessment program in reading. We are currently in the midst of the first full-scale year of reading assessment at Grade 4. As a matter of fact, today is the date on which some 9,000 or 10,000 Fourth Graders in Oregon are participating in that test administration.

TABLE 1. Number and Performance of Pupils, Reliability Estimates, and Correlations

SKILL AREA	N	Mean	S.D.	Reliability KR-20	CORRELATIONS Word Att.	Voc.	Comp.	Appl.
Word Attack	4098	31.1	6.9	.89	---	.64	.67	.70
Vocabulary	4098	6.8	1.8	.78	.64	---	.68	.63
Comprehension	4098	15.2	5.4	.87	.67	.68	---	.77
Application	4098	14.8	5.8	.88	.70	.63	.77	---

TABLE 2. Mean Scores Expressed as Percentages of Number of Items in Each Skill Area

SKILL AREA	No. of Items	Mean	Percentages
Work Attack	40	31.1	78
Vocabulary	8	6.8	85
Comprehension	21	15.2	72
Application	24	14.8	62

TABLE 3. Factor Analytic Structure of the Reading Test

READING OBJECTIVE	FACTOR 1	FACTOR 2	FACTOR 3
(Word Attack Objectives)			
1	.56		
2	.68		
3	.58		
4	.63		
5	.57	.32	
6	.68		
7	.54		
8	.54		
9	.62		
(Vocabulary Objectives)			
10	.41	.54	
11	.34	.57	
(Comprehension Objectives)			
12		.67	
13		.79	
14			.36
15		.62	.33
16		.75	
17		.66	
(Application Objectives)			
18	.30	.58	
19	.38	.52	
20	.31	.71	
21		.49	
22			.34
23			.75
24			.75
25			.52